NMR Structure of the N-terminal domain of the replication initiator protein DnaA

Thomas J. Lowery[†1,2], Jeffrey G. Pelton[†1], John-Marc Chandonia[1], Rosalind Kim[1], Hisao Yokota[1], & David E. Wemmer*[1,2]

[1]Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Berkeley, CA, 94720, USA.

[2]Department of Chemistry, University of California at Berkeley, Berkeley, CA, 94720. USA.

[†]These authors contributed equally to this work.

*Corresponding author:

David E. Wemmer
DEWemmer@lbl.gov
Phone: 510-486-4355
Fax: 510-486-6059

Running title: N-terminal domain of DnaA

DnaA is an essential component in the initiation of bacterial chromosomal replication. DnaA binds to a series of 9 base pair repeats leading to oligomerization, recruitment of the DnaBC helicase, and the assembly of the replication fork machinery. The structure of the N-terminal domain (residues 1 – 100) of DnaA from *Mycoplasma genitalium* was determined by NMR spectroscopy. The backbone r.m.s.d. for the first 86 residues was 0.6 +/- 0.2 Å based on 742 NOE, 50 hydrogen bond, 46 backbone angle, and 88 residual dipolar coupling restraints. Ultracentrifugation studies revealed that the domain is monomeric in solution. Features on the protein surface include a hydrophobic cleft flanked by several negative residues on one side, and positive residues on the other. A negatively charged ridge is present on the opposite face of the protein. These surfaces may be important sites of interaction with other proteins involved in the replication process. Together, the structure and NMR assignments should facilitate the design of new experiments to probe the protein-protein interactions essential for the initiation of DNA replication.

Keywords: replication, DnaB, DiaA, ATPase, initiation

The bacterial protein DnaA plays a central role in the initiation of chromosomal replication and is involved in the activation and repression of several genes, including its own gene (*dnaA*) in *E. coli*. Replication begins with the formation of oligomers of DnaA on recognition sites at the origin of replication (OriC). The DnaA complex unwinds and opens several AT-rich segments of double-stranded DNA within the origin and stabilizes the resulting single-strands. DnaA then recruits the DnaB helicase with help from DnaC to form the "prepriming complex" [1]. After priming of the single-stranded DNA, replication occurs by the action of DNA Polymerase III. ATP hydrolysis acts as the switch between active (ATP-bound) and inactive (ADP-bound) forms of DnaA. Initiation begins only with the ATP-bound form (for reviews see [2-4]).

The DnaA protein is composed of four functionally independent domains [4, 5]. Domains III and IV exhibit ATPase and DNA-binding activities, respectively, and are more highly conserved than the N-terminal domains I and II [3]. Evidence suggests that domains I and III are involved in oligomerization and binding DnaB, and that domain II is a flexible linker. Recent X-ray structures of domains III and IV of DnaA from *Aquifex aeolicus* bound to ADP [6] and a nonhydrolyzable ATP-analog [7], as well as domain IV bound to DNA [8] showed that domain III resembles other members of the AAA+ family of ATPases, and that domain IV adopts a helix-turn-helix DNA-binding fold. In contrast, the secondary structure of domain I has been predicted based on multiple sequence alignments [5, 9, 10], however, little is known about its tertiary structure.

DnaA recognizes five high affinity 9 base pair sites as well as several lower affinity 6 base pair sites within OriC via the C-terminal helix-turn-helix DNA-binding domain (domain IV) [11, 12]. In addition to direct binding, the formation of oligomers containing 20 to 30 molecules of DnaA at the origin has been demonstrated by electron microscopy [13, 14]. The mechanism of oligomer formation has been studied extensively [3, 15], but the nature of the complex remains poorly understood. Replacement of the dimerization domain of the lambda cI repressor with the N-terminal domain of DnaA resulted in a functional chimeric protein, suggesting direct interactions between the N-terminal domains [10, 16]. Results from solid phase binding assays [10] and gel mobility shift experiments [17]

have also suggested interactions between N-terminal domains. In the latter study, the N-terminal domain was shown to be required for cooperative binding of the protein to adjacent DnaA boxes. Mutational analysis has implicated Leu5, Trp6, Glu8, and Cys9 in domain I of *E. coli* DnaA as being critical for self-oligomerization [18, 19]. Other experiments have suggested a role for the ATPase domain in oligomerization [6, 7, 15, 16, 20]. Together, the observations that full-length DnaA from *E. coli* exists as a monomer in both the ADP- and ATP-bound forms, as well as in the presence of DNA that contains a single high affinity binding site [9, 12], have made it difficult to characterize the exact nature of self-oligomerization.

In addition to oligomerization, cross-linking studies have shown that the N-terminal domain of DnaA interacts directly with both DnaB [1] and the regulatory protein DiaA [21]. For DnaB, the regions of interaction were further defined by solid phase binding assays, which showed that residues 24-86 from the N-terminal domain, as well as residues in domain III, were important [20]. Within this set of residues, Pro28 appears to be especially important [19].

As part of a structural genomics effort we developed an algorithm to identify folded domains within larger proteins that had no known structural homologs [22]. The sequence corresponding to the N-terminal segment (residues 1 – 100) of DnaA from *Mycoplasma genitalium* was one such domain candidate. Initial NMR results indicated that this construct was indeed folded and a good candidate for structure determination. Herein, we present the structure of this N-terminal domain of DnaA, show that it exists as a monomer in solution, and relate it to previous genetic and biochemical data. The results presented provide a structural framework needed for the design of new experiments to further our understanding of the initiation of DNA replication.

Materials and Methods

The sequence containing domain I and a portion of domain II of DnaA from *M. genitalium* (gene MG469, gi number 12045329; residues 1 - 100) was selected and cloned as described [22, 23]. The protein was expressed as an N-terminal 6-His-TEV fusion protein in *E.coli* BL21* (DE3). In this

construct an N-terminal methionine and six glycine residues precede the first methionine of DnaA. These residues were not assigned or included in the structure calculations. Unlabeled protein was obtained by growing the cells in 1 L of LB at 37°C to an OD_{600} of 0.6 followed by induction with 1 mM isopropylthio-β-D-galactoside (IPTG) for 3.5 hrs. The molecular weight of the resulting protein was within one mass unit of the expected weight. Labeled protein was obtained by growing cells in 1 L of LB to an OD of 0.7, followed by pelleting the cells and transferring them to 0.75 L of isotope labeled M9 as described [24]. Isotopically labeled M9 contained either ^{15}N-labeled ammonium chloride or both ^{15}N-labeled ammonium chloride and ^{13}C-labeled glucose, to produce ^{15}N- and ^{15}N/^{13}C-labeled proteins, respectively. The cells were lysed in buffer containing 50 mM sodium phosphate (pH 7.5), 0.1 M sodium chloride, 1 mM PMSF, protease inhibitor cocktail (Roche Molecular Biochemicals, Mannheim, Germany), and cell debris was pelleted by ultracentrifugation. The soluble His-tagged protein was purified by nickel affinity chromatography using a running buffer of 50 mM sodium phosphate (pH 7.5), 0.1 M sodium chloride, 5% v/v glycerol, and 5 mM imidazole. After several column washes, the protein was eluted by increasing the imidazole concentration from 5 mM to 1 M. After concentration to approximately 2 mL by ultracentrifugation (5 kD MWCO Amicon Ultra), further purification was performed using size exclusion chromatography (HiLoad 16/60 Superdex 75pg; GE Healthcare) with a running buffer containing 50 mM sodium phosphate (pH 7.5) and 0.15 M sodium chloride. Fractions containing the N-terminal domain were pooled and treated with TEV protease overnight at room temperature. The protein was then reapplied to a nickel-affinity column and the eluate was collected and concentrated. Analytical size exclusion chromatography was conducted on a BioCad Sprint II with a TSK-Gel G4000SW_{xl} column (Tosoh Bioscience LLC; 7.8 mm ID x 30 cm) using bovine serum albumin, carbonic anhydrase, and lysozyme as molecular weight standards. Analytical ultracentrifugation (AUC) was performed using a Beckman Optima XL-I analytical ultracentrifuge. Purified protein was dialyzed against 50 mM sodium phosphate (pH 6.8) and 0.1 M sodium chloride overnight at 4°C. Protein at a concentration of 140 μM was loaded into each of five sample cells (100

μL) as determined by UV absorbance (ε_{280} = 5.12 mM^{-1} cm^{-1}). Blank cells were filled with dialysis buffer (110 μL). The centrifuge was allowed to equilibrate for 15 hours at 30,000 rpm and 20°C before the recording of 10 sequential scans. Data were fit using UltraScan software [25]. NMR samples were prepared by concentrating the purified protein in buffer containing 50 mM sodium phosphate (pH 6.8), 0.1 M sodium chloride, 2 mM EDTA, 0.05% azide, protease inhibitor cocktail (Roche Molecular Biochemicals, Mannheim, Germany), and 0.1 mM PMSF to a concentration of 1.0 to 1.5 mM. NMR samples in D_2O were prepared by lyophilizing an existing NMR sample and re-dissolving it into D_2O. NMR spectra were recorded on Bruker NMR spectrometers, including a DRX 500 equipped with a cryoprobe, a DMX 600, an Avance 800 equipped with a cryoprobe, and an Avance II 900. All spectra were recorded at 25°C. Backbone resonances were assigned with 3D HNCA [26] , HNCO [27, 28], CBCA(CO)NH [29], HNCACB [30] spectra recorded at 500 MHz and HBHACONH [31] and HNCACO [27] spectra recorded at 600 MHz. Side-chain aliphatic $^{13}C/^{1}H$ signals were assigned with 3D-HBHA(CO)NH, 3D-(H)C(CO)NH [32], 3D-H(CCO)NH [32], 3D-^{15}N TOCSY [33], and 3D-HCCH-TOCSY [34] spectra at 600 MHz. In the final analysis, all of the backbone ^{1}H,^{15}N,$^{13}C\alpha$, and $^{1}H\alpha$ signals and 88% of the sidechain signals were assigned, excluding the N-terminal methionine and six glycine residues that result from cloning, as well as the first methionine residue (Met1), Phi torsion angle restraints were obtained from a 3D HNHA spectrum [35] recorded at 800 MHz. Stereospecific assignments for the methyl groups of 3 of 4 valine and 11 of 12 leucine residues were obtained by comparison of HSQC spectra of fully and 10% ^{13}C-labeled samples [36, 37]. NOEs were identified in 3D ^{15}N-^{1}H NOESY-HSQC (800 MHz), 3D ^{13}C-^{1}H NOESY-HSQC (800 and 900 MHz), and a 4D ^{13}C-HMQC-NOESY-HSQC [38] (600 MHz) spectra with mixing times ranging between 85 and 100 ms. A total of 53 ^{1}H-^{15}N and and 35 ^{1}H-$^{13}C\alpha$ residual dipolar couplings (rdcs) were used in the structure calculations. The ^{1}H-^{15}N rdcs were obtained from IPAP [39] spectra recorded on a uniformly ^{15}N-labeled protein sample, and the ^{1}H-$^{13}C\alpha$ dipolar couplings were based on a 3D (HA)CA(CO)NH spectrum [40], recorded on a uniformly - $^{15}N/^{13}C$ labeled protein sample. Both proteins were dissolved

in 12 mg/mL pf1 phage [41]. Tensor parameters were determined from histogram plots of the couplings [42] and values based on intermediate structures [43, 44]. The magnitude of the alignment tensor and rhombicity were set to -13.7 Hz and 0.47 for HN vectors and -13.3 Hz and 0.47 for CαH vectors. Force constants were adjusted to 0.4 kcal $mol^{-1}Hz^{-2}$ and 0.2 kcal mol^{-1} Hz^{-2} to yield r.m.s.d.s equal to the uncertainties in the measurements (Table 1). Structures were calculated using a torsion-angle simulated annealing protocol [45, 46] as implemented in the XPLOR-NIH program, version 2.11.2 [47]. Center averaging and distance corrections [48] were used for non-stereoassigned protons. The NOE and rdc data were generally consistent with one another as evidenced by the fact that average structures (20 best of 200) calculated with and without rdcs superimposed with an r.m.s.d of 0.67 Å. The main difference was a slight displacement of the end of helix III at Asn53. The number of restraints and structure statistics are presented in Table 1.

Results and Discussion

As discussed above, the results from several studies suggested that the N-terminal domain dimerizes or assists in the formation of larger oligomers [10, 18, 19]. Before initiating NMR studies, we probed the oligomerization state of our construct using both size exclusion chromatography (SEC) and analytical ultracentrifugation (AUC). The initial SEC results suggested that the protein was a dimer at a concentration of 1 mM (data not shown). In contrast, the AUC results, which were conducted at a concentration of 140 μM, revealed that the protein was monomeric (data not shown). In addition, identical ^{1}H-^{15}N HSQC spectra were obtained at 70 μM and the working concentration of 1.2 mM (Figure 1), indicating that protein exists in the same state at both concentrations. Because results from AUC depend only on the mass of the protein, whereas results from SEC are influenced by the shape of the molecule, we conclude that the protein is a monomer in solution at millimolar concentrations, and that the apparent molecular weight determined from SEC was an artifact, likely resulting from the first 6 and last 15 residues being unstructured and flexible in solution (see below).

A summary of the NMR-derived restraints and structural statistics are presented in Table 1. The NMR data reveal that domain I is composed of five helices and a three-stranded β-sheet with topology ααββααβα (Figure 2). The experimentally determined secondary structure is in general agreement with predictions based on multiple sequence alignments [5, 9, 10]. The exceptions are the presence of β-strand 1 (residues 31-34), a break between helices III and IV, and the presence of helix V (residues 77 to 84). Helix V may not be present in *E. coli* since there are two proline residues (Pro82 and Pro85) in this region. The three-dimensional structure is formed by the packing of helices I and IV and II and III against one another and one side of the β-sheet, while helix V packs against the opposite face (Figure 3). As mentioned previously, the last fifteen residues of the construct, which correspond to the first residues of domain II [4], show no long-ranges NOEs, suggesting that they are unstructured. These data are consistent with this segment functioning as a flexible linker between domain I and the ATPase and HTH domains [4]. A search for structural homologs [49] revealed that the N-terminal domain is similar to ribosomal binding factor A (RbfA) [50, 51], which contains a type II KH RNA-binding fold (Z score of 3-4). However, the structural similarity probably does not carry over to function. To our knowledge, there is no evidence that the N-terminal domain of DnaA binds to nucleic acids.

A total of 17 residues are conserved among the roughly 212 known sequences of the N-terminal domain of DnaA (Figure 2). Approximately 80% of these residues are involved in hydrophobic packing within the protein core (Figure 3b). Three other conserved residues, Asn45, Asn53, and Glu57, are located on the surface of the protein in helices III and IV. Comparison also shows that proline residues in *E. coli* and the consensus sequence (not shown) map to loops and unstructured regions of the structure (Figure 2). The only exceptions are the two proline residues in *E. coli* that map to helix V, as mentioned above. Similarly, the locations of insertions in other sequences map to loops or unstructured regions. For insertions, the only exceptions occur in three mycobacterium sequences, in which there is

an insertion in helix II. Together, these observations indicate that the overall fold of the N-terminal domain is conserved across all species.

As was previously noted [5, 9, 10], the sequence similarity within domain I is lower than within domains III and IV, making it difficult to identify functionally important residues. Experimentally, residues at positions Gln3, Phe4, and Phe7 (Leu5, Trp6 and Cys9 in *E.coli*) have been shown to be critical for oligomerization [18, 19] and residues 24-86, and in particular Pro28, have been implicated in DnaB binding [19, 20]. Residues Gln3, Phe4, and Phe7 line the inner face of helix I (Figure 3) and Pro28 is in the loop connecting helix II and β-strand I. These secondary structure elements form a hydrophobic cleft with positive residues along the bottom edge (Gln3, Lys8, Lys12, and Lys13), and negative residues along the top edge (Asp23 and Asp28) (Figure 3c). Approximately half of the residues lining this cleft are conserved in the full multiple sequence alignment. On the opposite face of the protein, a negatively charged ridge runs from Glu46 to Glu67 along the axes of helices III and IV (Figure 3c). These surface features may be important in the formation of DnaA oligomers and binding with DnaB, DiaA, and other proteins involved in the replication process.

DnaA oligomerizes at the origin of replication and interacts with a host of other proteins, including DnaB and DiaA, during DNA replication [2, 3]. To date, the nature of these interactions and the specific role of the N-terminal domain remain poorly understood. The chemical shift assignments presented here should promote efforts to explore these interactions directly at the molecular level. In addition, the structure and surface analysis should enable the design of more precise genetic studies, which, together with binding studies, should help unravel the complex set of protein-protein interactions that are central to this fundamental process.

Assession codes

NMR restraints and coordinates have been deposited in the Protein Data Bank under accession number 2JMP. Chemical shift assignments have been deposited in the BioMagResBank (BMRB) under accession number BMRB-15055.

Acknowledgements

We thank D. King for mass spectral analysis, and J. Erzberger, M. Mott, and J. Berger for helpful discussions. We also thank M. Doucelff for recording several spectra. T.J.L. acknowledges the University of California Biotechnology Research and Education Program for financial support. This work was supported by grant GM62412 to the Berkeley Structural Genomics Center from the National Institute of General Medical Sciences, National Institutes of Health. We also thank NSF (BBS87W134 and 0119304) and NIH (RR15756 and GM068933) for funding NMR equipment.

Figure legends

Figure 1. ^{1}H-^{15}N HSQC spectra recorded at 500 MHz of the N-terminal domain of DnaA from *M. genitalium* at concentrations of (a) 70 μM and (b) 1.2 mM.

Figure 2. Sequences of the N-terminal domains of DnaA from *M. genitalium* and *E. coli*. The alignment is based on 212 DnaA sequences. The domain boundaries are taken from Messer et al. 1999[4]. Domain III begins at residue 134 in *E. coli*. Helices and β-strands found in DnaA from *M. genitalium* are shown schematically above the sequences as black bars and arrows, respectively. Asterisks (*) denote the positions of conserved residues within the multiple sequence alignment, carets (^) denote the locations of proline residues in *E. coli* or the consensus sequence, and bars (|) indicate the positions of insertions in other DnaA sequences. Hydrophobic residues common in both *M. genitalium* and *E. coli* are shaded grey.

Figure 3. (a) Backbone trace for the top 20 structures superimposed using the backbone N, Cα, and C' coordinates. Unstructured residues (89-100) are not shown. (b) Ribbon diagram of the structure. The locations of conserved hydrophobic residues based on the full multiple sequence alignment are shown in yellow. Sites of functionally critical residues are denoted with black arrows. (c) Electrostatic surface of the structure showing positively charged regions in blue, neutral regions in white, and negatively charged regions in red. (c - left) The hydrophobic cleft near helix I is apparent, with positive residues along the bottom edge and negative residues along the top edge (see text). (c - right) A 180 degree rotation of the molecule shows the negatively charged surface associated with helices III and IV. Figures 3b and 3c were made with PyMOL (Delano Scientific, Palo Alto, Ca.).

Table 1. Structural statistics for the N-terminal domain of DnaA from *M. genitalium.*

Parameter	20 Conformers[a]
RMSD from restraints[b]	
NOE (Å)	0.014 (±0.002)
Cdih (°)	0.37 (±0.09)
HN rdcs R-factor (%) / rmsd (Hz)	2.5 (±0.3%) / 0.46 (±0.06)
$C^{\alpha}H$ rdcs R-factor (%) / rmsd (Hz)	9.5 (±0.7%) / 1.7 (±0.1)
RMSD from idealized geometry	
Bond (Å)	0.00174 (±0.00007)
Angle (°)	0.426 (±0.006)
Improper (°)	0.25 (±0.008)
Number of violations / structure	
NOE (> 0.1 Å)	1.9 (±1.5)
Cdih (> 5 °)	0
HN rdcs (> 1.5 Hz)	0.3 (±0.07)
$C^{\alpha}H$ RDCs (> 3 Hz)	3 (±1)
Force field energy[c]	-1100 (±200) kJ/mol
Coordinate RMSDs to the mean[d]	
Secondary structure / backbone (Å)	0.6 (±0.2)
Residues 3 - 86 / backbone (Å)	0.7 (±2)
Heavy atom < 40% solvent exposed (Å)	1.2 (±0.1)
Procheck Analysis[e]	
Most favored regions (%) secondary structure (all)	80.9 (58)
Additional allowed regions (%)	19.1 (35.9)
Generously allowed regions (%)	0 (4.5)
Disallowed regions (%)	0 (1.1)

[a]Statistics are presented for the best 20 of 200 structures.

[b]A total of 286 intra-residue, 194 sequential, 103 medium-range, and 159 long-range NOEs, 50 hydrogen bonds (2 per hydrogen bond), 46 phi angle restraints, 53 HN residual dipolar couplings (rdcs), and 35 CαH rdcs were used in the calculations. The RDC R-factor is defined as the ratio of the r.m.s. deviation between observed and calculated values to the expected r.m.s. deviation if the vectors were randomly distributed[52].

[c] The negative Lennard-Jones energy indicates no large van der Waals clashes. Determined using the GROMOS96 implementation within the Swiss-PDBViewer [53].

[d]Secondary structure included N,Cα,C' atoms for helical regions 3-14,19-24,46-52,59-62,77-85, and β-strand regions 32-34,39-42,72-75. Less than 40% solvent exposed residues included 4,6-11,14-16,18-23,25-29,31-34,36-43,45,48,49,51-53,55-59,61-66,68,69,71-77,79-81,83-85.

[e]Version 3.5.4 [54]

References

1. Marszalek, J., and Kaguni, J. M. (1994) *J. Biol. Chem.* **269,** 2883-2890.
2. Kato, J. (2005) *Crit. Rev. Biochem. Mol. Biol.* **40,** 331-342.
3. Messer, W. (2002) *FEMS Microbiol. Rev.* **26,** 355-374.
4. Messer, W., Blaesing, F., Majika, J., Nardmann, J., Schaper, S., Schmidt, A., Seitz, H., Speck, C., Tuengler, D., Wegrzyn, G., Weigel, C., Welzeck, M., and Zakrzewska-Czerwinska, J. (1999) *Biochemie* **81,** 819-825.
5. Sutton, M. D., and Kaguni, J. M. (1997) *J. Mol. Biol.* **274,** 546-561.
6. Erzberger, J. P., Pirruccello, M. M., and Berger, J. M. (2002) *EMBO J.* **21,** 4763-4773.
7. Erzberger, J. P., Mott, M. I., and Berger, J. M. (2006) *Nat. Struct. Biol.* **13,** 676-683.
8. Fujikawa, N., Kurumizaka, H., Nureki, O., Terada, T., Shirouzu, M., Katayama, T., and Yokoyama, S. (2003) *Nucleic Acids Res.* **31,** 2077-2086.
9. Schaper, S., and Messer, W. (1995) *Proteins* **28,** 1-9.
10. Weigel, C., Schmidt, A., Seitz, H., Tuengler, D., Welzeck, M., and Messer, W. (1999) *Mol. Microbiol.* **34,** 53-66.
11. Speck, C., and Messer, W. (2001) *EMBO J.* **20,** 1469-1476.
12. Speck, C., Weigel, C., and Messer, W. (1999) *EMBO J.* **18,** 6169-6176.
13. Crooke, E., Thresher, R., Hwang, D. S., Griffin, J., and Kornberg, A. (1993) *J. Mol. Biol.* **233,** 16-24.
14. Funnell, B. E., Baker, T. A., and Kornberg, A. (1987) *J. Biol. Chem.* **262,** 10327-10334.
15. Messer, W., Blaesing, F., Jakimowicz, D., Krause, M., Majka, J., Nardmann, J., Schaper, S., Seitz, H., Speck, C., Weigel, C., Wegrzyn, G., Welzeck, M., and Zakrzewska-Czerwinska, J. (2001) *Biochimie* **83,** 5-12.

16. Jakimowicz, D., Majka, J., Konopu, G., Wegrzyn, G., Messer, W., Schrempf, H., and Zakrzewska-Czerwinska, J. (2000) *J. Mol. Biol.* **298,** 351-364.

17. Majka, J., Zakrzewska-Czerwinska, J., and Messer, W. (2001) *J. Biol. Chem.* **276,** 6243-6252.

18. Felczak, M., Simmons, L. A., and Kaguni, J. M. (2005) *J. Biol. Chem.* **280,** 24627-24633.

19. Simmons, L. A., Felczak, M., and Kaguni, J. M. (2003) *Mol. Microbiol.* **49,** 849-858.

20. Seitz, H., Weigel, C., and Messer, W. (2000) *Mol. Microbiol.* **37,** 1270-1279.

21. Ishida, T., Akimitsu, N., Kashioka, T., Hatano, M., Kubota, T., Ogata, Y., Sekimizu, K., and Katayama, T. (2004) *J. Biol. Chem* **279,** 45546-45555.

22. Chandonia, J. M., Kim, S. H., and Brenner, S. E. (2006) *Proteins* **62,** 356-370.

23. Shin, D. H., Kim, J. S., Yokota, H., Kim, R., and Kim, S. H. (2006) *Protein Sci.* **15,** 921-928.

24. Marley, J., Lu, M., and Bracken, C. (2001) *J. Biomol. NMR* **20,** 71-75.

25. Demeler, G. (2005) *in* "Modern Analytical Ultracentrifugation: Techniques and Methods" (Scott, D. J., and Rowe, A. J., Eds.), pp. 210-229, Royal Society of Chemistry, UK.

26. Grzeisek, S., and Bax, A. (1992) *J. Magn. Reson.* **99,** 201-207.

27. Kay, L. E., Xu, G. Y., and Yamazaki, T. (1994) *J. Magn. Reson.* **109 A,** 129-133.

28. Stonehouse, J., Clowes, R. T., Shaw, G. L., Keeler, J., and Laue, E. D. (1995) *J. Biomol. NMR* **5,** 226-232.

29. Grzeisek, S., and Bax, A. (1992) *J. Am. Chem. Soc.* **114,** 6291-6293.

30. Wittekind, M., and Mueller, L. (1993) *J. Magn. Reson.* **101 B,** 201-205.

31. Grzesiek, S., and Bax, A. (1993) *J. Biomol. NMR* **3,** 185-204.

32. Grzesiek, S., Anglister, J., and Bax, A. (1993) *J. Magn. Reson.* **101 B,** 114-119.

33. Driscoll, P. C., Clore, G. M., Marion, D., Wingfield, P. T., and Gronenborn, A. M. (1990) *Biochemistry* **29,** 3542-3556.

34. Kay, L. E., Xu, G. Y., Singer, A., Muhandram, D., and Forman-Kay, J. (1993) *J. Magn. Reson.* **101 B,** 333-337.

35. Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) *J. Biomol. NMR.* **4,** 871-878.

36. Neri, D., Szyperski, T., Otting, G., Senn, H., and Wuthrich, K. (1989) *Biochemistry* **28,** 7510-7516.

37. Szyperski, T., Neri, D., Leiting, B., Otting, G., and Wuthrich, K. (1992) *J. Biomol. NMR* **2,** 323-334.

38. Vuister, G. W., Clore, G. M., Gronenborn, A. M., Powers, R., Garrett, D. S., Tschudin, R., and Bax, A. (1993) *J. Magn. Reson.* **101 B,** 210-213.

39. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson.* **131,** 373-378.

40. Tjandra, N., and Bax, A. (1997) *J. Am. Chem. Soc.* **119,** 9576-9577.

41. Hansen, M. R., Mueller, L., and Pardi, A. (1998) *Nat. Struct. Biol.* **5,** 1065-1074.

42. Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) *J. Magn. Reson.* **133,** 216-221.

43. Losonczi, J. A., Andrec, M., Fischer, M. W., and Prestegard, J. H. (1999) *J. Magn. Reson.* **138,** 334-342.

44. Zweckstetter, M., and Bax, A. (2000) *J. Am. Chem. Soc.* **122,** 3791-3792.

45. Rice, L. M., and Brunger, A. T. (1994) *Proteins* **19,** 277-290.

46. Stein, E. G., Rice, L. M., and Brunger, A. T. (1997) *J. Magn. Reson.* **124,** 154-164.

47. Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) *J. Magn. Reson.* **160,** 65-73.

48. Fletcher, C. M., Jones, D. N. M., Diamond, R., and Neuhaus, D. (1996) *J. Biomol. NMR* **8,** 292-310.

49. Holm, L., and Sander, C. (1993) *J. Mol. Biol.* **233,** 123-138.

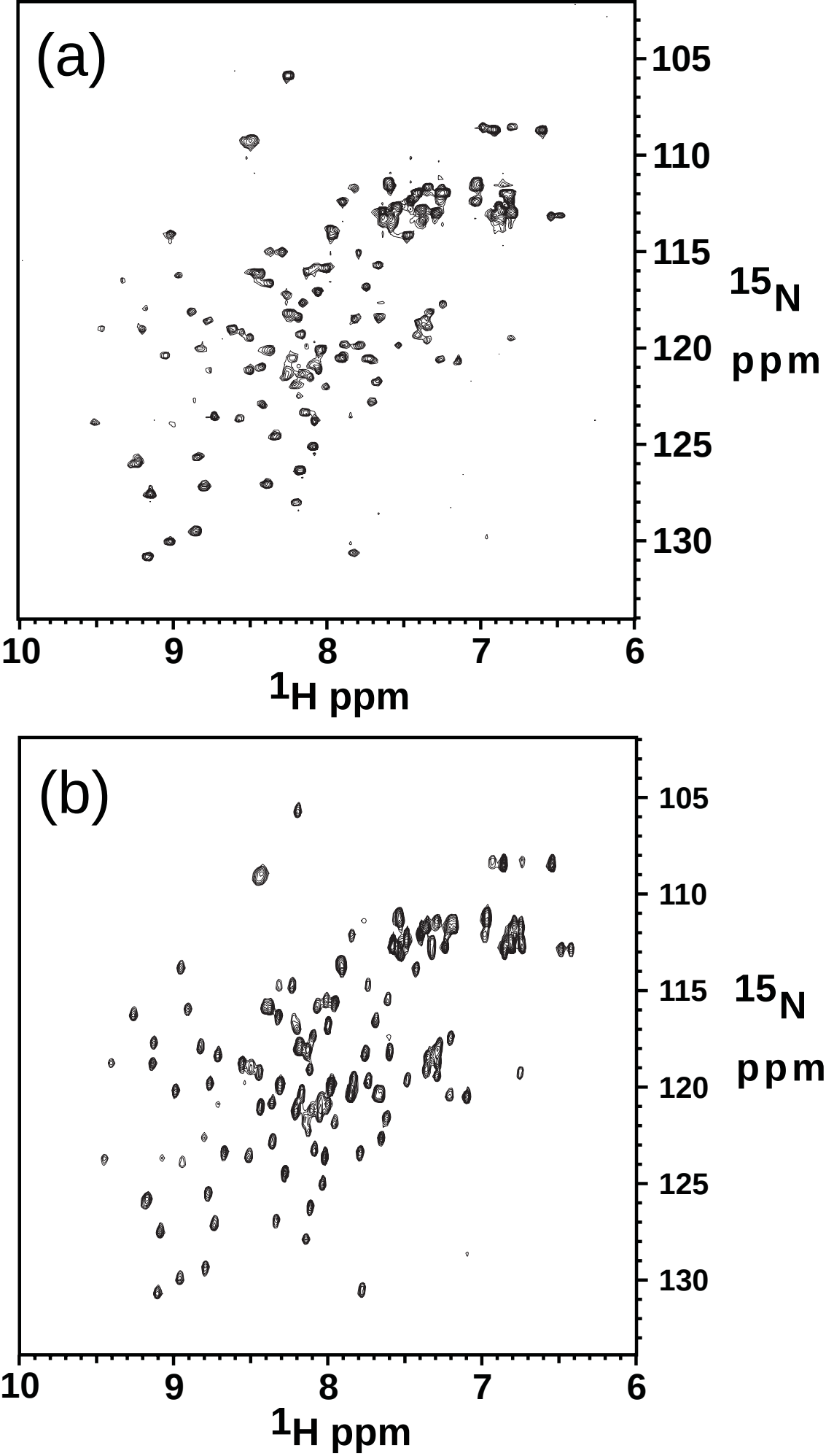
50. Huang, Y. J., Swapna, G. V., Rajan, P. K., Ke, H., Xia, B., Shukla, K., Inouye, M., and Montelione, G. T. (2003) *J. Mol. Biol.* **327,** 521-536.

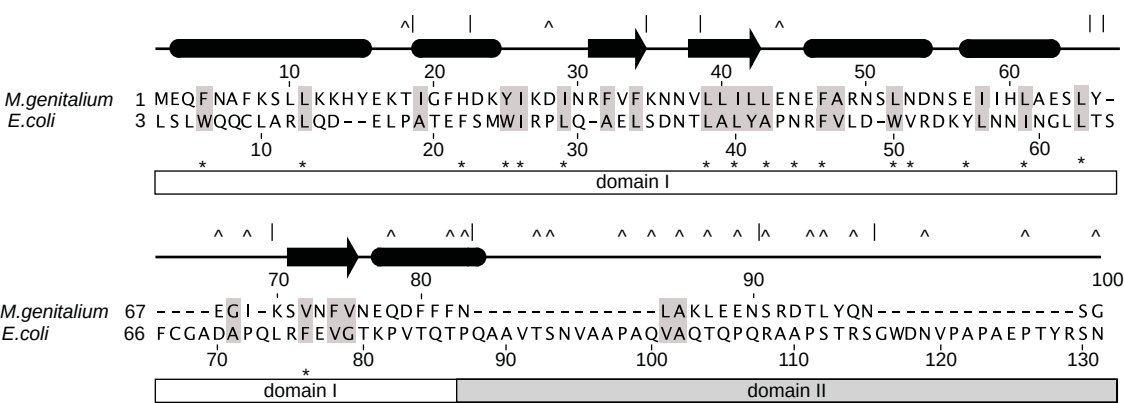
51. Rubin, S. M., Pelton, J. G., Yokota, H., Kim, R., and Wemmer, D. E. (2003) *J. Struct. Funct. Genomics* **4,** 235-243.

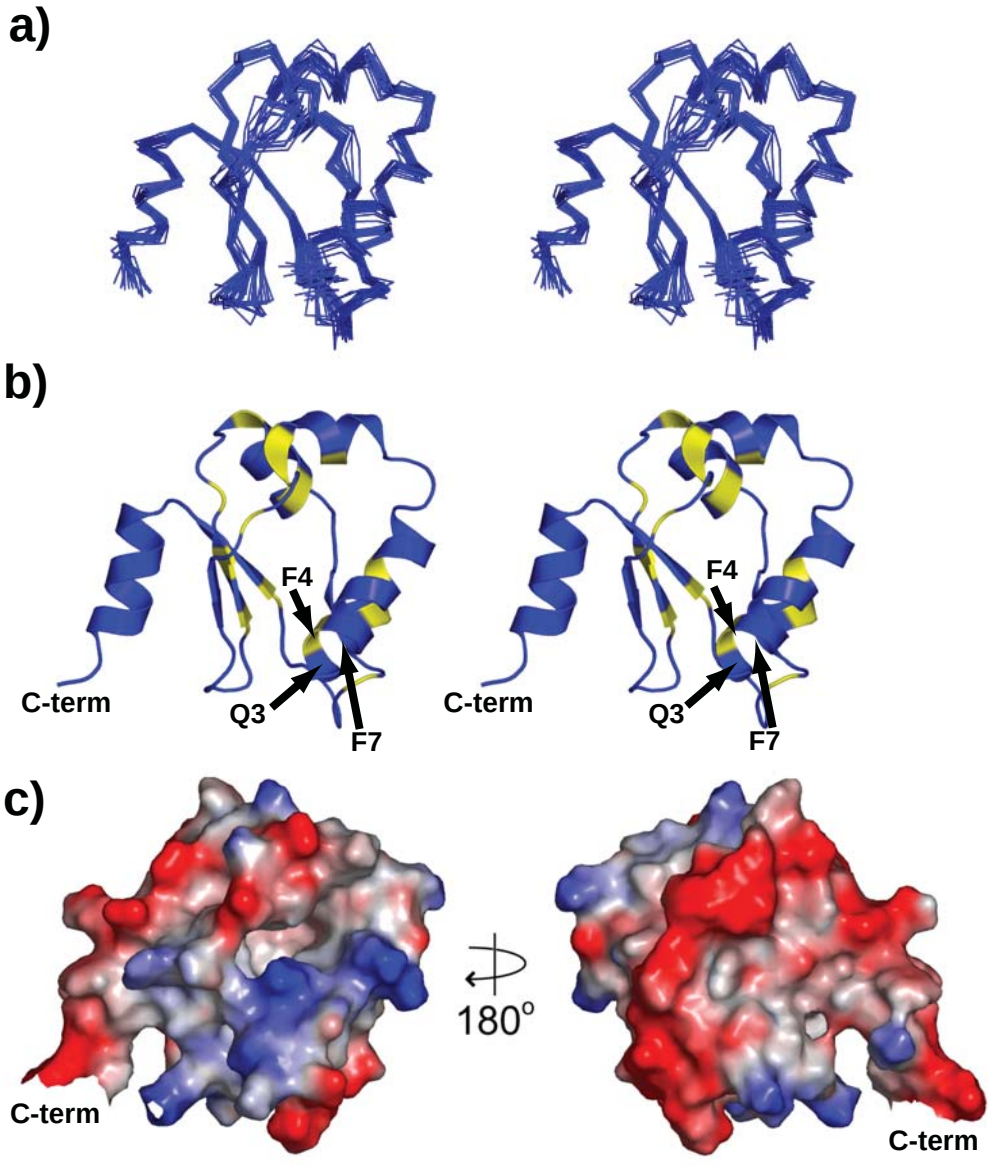
52. Bax, A., Kontaxis, G., and Tjandra, N. (2001) *Methods Enzymol.* **339,** 127-174.

53. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18,** 2714-2723.

54. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR.* **8,** 477-486.


(a)
105
110
115
120
125
130
15N ppm
10
9
8
7
6
1H ppm
(b)
105
110
115
120
125
130
15N ppm
10
9
8
7
6
1H ppm


M.genitalium 1 MEQFNAFKSLLKKHYEKTIGFHDKYIKDINRFVFKNNVLLILLENEFARNSLNDNSEIIHLAESLY-
E.coli 3 LSLWQQCLARLQD--ELPATEFSMWIRPLQ-AELSDNTLALYAPNRFVLD-WVRDKYLNNINGLLTS
domain I
M.genitalium 67 ----EGI-KSVNFVNEQDFFFN----------LAKLEENSRDTLYQN----------SG
E.coli 66 FCGADAPQLRFEVGTKPVTQTPQAAVTSNVAAPAQVAQTQPQRAAPSTRSGWDNVPAPAEPTYRSN
domain I
domain II


a)
b)
F4
Q3
F7
C-term
F4
Q3
F7
C-term
c)
C-term
180°
C-term